(18)F-FDG PET image biomarkers improve prediction of late radiation-induced xerostomia

van Dijk, Lisanne V; Noordzij, Walter; Brouwer, Charlotte L; Boellaard, Ronald; Burgerhof, Johannes G M; Langendijk, Johannes A; Sijtsema, Nanna M; Steenbakkers, Roel J H M

Published in:
Radiotherapy and Oncology

DOI:
10.1016/j.radonc.2017.08.024

Publication date:
2018

Radiotherapy and Oncology 126 (2018) 89–95

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Xerostomia and image biomarkers

^{18}F-FDG PET image biomarkers improve prediction of late radiation-induced xerostomia

Lisanne V. van Dijk [a,*], Walter Noordzij [b], Charlotte L. Brouwer [a], Ronald Boellaard [b], Johannes G.M. Burgerhof [c], Johannes A. Langendijk [a], Nanna M. Sijtsema [a], Roel J.H.M. Steenbakkers [a]

[a] Department of Radiation Oncology; [b] Nuclear Medicine and Molecular Imaging; and [c] Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 20 March 2017
Received in revised form 31 July 2017
Accepted 21 August 2017
Available online 23 September 2017

Keywords:
Xerostomia
NTCP
Image biomarkers
Head and neck cancer
FDG-PET
Radiomics

ABSTRACT

Background and purpose: Current prediction of radiation-induced xerostomia 12 months after radiotherapy (Xer_{12m}) is based on mean parotid gland dose and baseline xerostomia ($Xer_{baseline}$) scores. The hypothesis of this study was that prediction of Xer_{12m} is improved with patient-specific characteristics extracted from ^{18}F-FDG PET images, quantified in PET image biomarkers (PET-IBMs).

Patients and methods: Intensity and textural PET-IBMs of the parotid gland were collected from pre-treatment ^{18}F-FDG PET images of 161 head and neck cancer patients. Patient-rated toxicity was prospectively collected. Multivariable logistic regression models resulting from step-wise forward selection and Lasso regularisation were internally validated by bootstrapping. The reference model with parotid gland dose and $Xer_{baseline}$ was compared with the resulting PET-IBM models.

Results: High values of the intensity PET-IBM (90th percentile (P90)) and textural PET-IBM (Long Run High Grey-level Emphasis 3 (LRHG3E)) were significantly associated with lower risk of Xer_{12m}. Both PET-IBMs significantly added in the prediction of Xer_{12m} to the reference model. The AUC increased from 0.73 (0.65–0.81) (reference model) to 0.77 (0.70–0.84) (P90) and 0.77 (0.69–0.84) (LRHG3E).

Conclusion: Prediction of Xer_{12m} was significantly improved with pre-treatment PET-IBMs, indicating that high metabolic parotid gland activity is associated with lower risk of developing late xerostomia. This study highlights the potential of incorporating patient-specific PET-derived functional characteristics into NTCP model development.

^{18}F-FDG PET imaging provides functional information about the metabolic activity of tissue. This makes ^{18}F-FDG PET a powerful and widely used diagnostic modality in oncology. In head and neck oncology, ^{18}F-FDG PET can complement other image modalities in tumour staging and delineation for radiotherapy [1,2]. The common clinical use of ^{18}F-FDG PET allows for the possibility to extract large amounts of patient-specific functional information that could contribute to prognosis for head and neck cancer (HNC) patients. Several studies have shown that PET image characteristics of the tumour can contribute to predicting overall, disease-free or event-free survival [3–6]. However, patient-specific image characteristics for predicting normal tissue radiation toxicities are less explored, while these are also crucial in supporting treatment decisions. Additionally, new radiation techniques (e.g. proton therapy [7] and magnetic resonance imaging (MRI) guided radiation [8]) may allow for better sparing of normal tissue. These new techniques demand improved prediction models, to select patients most at risk of developing toxicities [9].

Radiation-induced xerostomia is a major and frequent side effect for HNC patients, and has a considerable impact on these patients' quality of life [10]. Conventional Normal Tissue Complication Probability (NTCP) models that predict patient-rated xerostomia are based on dose–volume parameters and baseline complaints [11,12]. However, there is still a significant, unexplained variance in predicting xerostomia with these models. Therefore, the demand persists to improve the identification of patients at risk. Previous work showed that patient-specific CT characteristics of the parotid glands could significantly improve the prediction of patient-rated xerostomia, however, model performance improvement was marginal [13]. The hypothesis was that the predictive CT characteristic is related to the ratio of non-function to functional parotid tissue. It can be expected that this ratio would be better represented by image characteristics from functional imaging (i.e. PET or MR images).

* Corresponding author at: Department of Radiation Oncology, University Medical Center Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands.
E-mail address: l.v.van.dijk@umcg.nl (L.V. van Dijk).

https://doi.org/10.1016/j.radonc.2017.08.024
0167-8140/

In this study, the relationship was tested between metabolic activity of the parotid gland and late xerostomia. Consequently, the patient-specific response to radiation in developing this toxicity was investigated. The purpose was to determine whether functional information from ^{18}F-FDG PET images, which is quantified in PET-image biomarkers (PET-IBMs), was associated with patient-rated moderate-to-severe xerostomia 12 months after radiotherapy (Xer_{12m}). Since current NTCP prediction models are based on parotid gland dose and baseline complaints, the study subsequently addressed whether PET-IBMs could improve on the current prediction of Xer_{12m}

Materials and methods

Patient demographics and treatment

^{18}F-FDG PET/CT scans were acquired of 161 HNC patients in treatment position before the start of radiotherapy. The patients were treated with definitive radiotherapy either with or without concurrent chemotherapy or cetuximab, between November 2010 and August 2015. Patients without follow-up data 12 months after radiotherapy were excluded from this study. Patients were also excluded if they underwent surgery in the head and neck area before or within one year after treatment.

A detailed description of the radiotherapy protocols is given in previous studies [13,14]. In summary, all patients were treated with IMRT or VMAT using a simultaneous integrated boost (SIB) technique. The parotid glands and the swallowing structures were spared as much as possible without compromising the dose to the target volumes [14,15]. Patients received a total dose of 70 Gy (2 Gy per fraction, 5 or 6 times a week) to the primary tumour and, if present, pathological lymph nodes. A radiation dose of 54.25 Gy (1.55 Gy per fraction, 5 or 6 times a week) was delivered to the elective lymph node levels.

Endpoints

The primary endpoint was patient-rated moderate-to-severe xerostomia 12 months after radiotherapy (Xer_{12m}), which corresponds to the 2 highest scores of the 4-point Likert scale of the EORTC QLQ-H&N35 questionnaire. This endpoint was prospectively assessed as part of a Standard Follow-up Program (SFP) for Head and Neck Cancer Patients (NCT02435576), as described in previous studies [11,12,16].

Dose and clinical parameters

For treatment planning, parotid glands were delineated on the planning (PET/)CT scans. The mean dose to both the contra- and ipsilateral parotid and submandibular glands were extracted from the dose–volume information [11,17]. In addition, baseline patient-rated xerostomia ($Xer_{baseline}$) was also considered (none vs. any).

Patient characteristics such as age, sex, WHO-performance, tumour stage and body mass index did not significantly add to the parotid gland dose and $Xer_{baseline}$ in predicting Xer_{12m} in previous studies [11,13,18]. This was again observed in the current cohort, therefore these variables were not further reported in this study.

^{18}F-FDG PET acquisition

Approximately 2 weeks before the start of radiotherapy, ^{18}F-FDG PET/CT images (Siemens Biograph 64-slice PET/CT scanner, Siemens Medical Systems, Knoxville, TN, USA) were acquired in with the patient positioned for radiotherapy. PET/CT system performance were initially harmonised conform the Netherlands protocol for FDG PET imaging [19] and later by EARL accreditation [20].

Patients were instructed not to eat or drink 6 h before scanning, but were encouraged to drink water to ensure adequate hydration. A body weight-based intravenous injection dose of 3 MBq/kg was administered 60 min prior to the ^{18}F-FDG PET acquisition. ^{18}F-FDG PET images were acquired in the caudal–cranial direction with an acquisition time of ~3 min per bed position.

Candidate PET-image biomarkers

Intensity PET-IBMs were extracted, representing first order standardised uptake value (SUV) characteristics of the delineated contra-lateral parotid glands. Examples are mean, minimum, maximum, standard deviation and root mean square of the SUVs. For the complete list of the 24 intensity PET-IBMs, see Supplementary data 1. Fig. 1 shows a schematic representation of PET-IBMs' extraction process.

Furthermore, more complex, textural features were extracted describing the intensity heterogeneity. These textural PET-IBMs were extracted from the grey level co-occurrence matrix (GLCM) [21], grey level run-length matrix (GLRLM) [22,23], grey level size-zone matrix (GLSZM) [24] and neighbourhood grey tone difference matrix (NGTDM) [25]. GLCM describes the grey level transitions. GLRLM and GLSZM describe the directional and volumetric grey level repetitions, respectively. NGTDM describes the relationship of sum and averages of grey level differences of direct adjacent voxels.

For this study, the average of PET-IBMs from GLCM and GLRLM in 13 independent directions was used. The range of SUVs was binned with a fixed bin size of 0.25. Discretisation of SUV is necessary to reduce the number of possible intensity values, and so reduce noise when calculating textural features [26]. All 66 textural PET-IBMs (25 GLCM, 18 GLRLM, 18 GLSZM and 5 NGTDM) were normalised by subtracting the average from the PET-IBMs' values and then dividing by the standard deviation. For the complete list refer to Supplementary data 2. All PET-IBMs were extracted in MATLAB (version R2014a).

Univariable analysis

Univariable logistic regression analysis was performed to evaluate the basic associations of PET-IBMs with late xerostomia. p-Values <0.05 were considered statistically significant. Coefficients (β) were evaluated to understand the effect that is described by the PET-IBMs in relation to Xer_{12m}. The univariable analysis was not used for the variable selection.

Multivariable analysis

Reference model

A reference prediction model was evaluated for the current patient cohort. This model was based on the mean dose to the contralateral parotid gland and $Xer_{baseline}$. These were the predictors that were identified by Beetz et al. [11].

Intensity and textural PET-IBMs

First, a basic PET-IBM model was created by adding the 'mean SUV' of the parotid gland as an extra variable to the reference model. Since this variable is the simplest of PET-IBMs, it is the easiest to interpret.

Both step-wise forward selection and Lasso regularisation were performed for multivariable logistic analysis of the PET-IBMs, together with parotid dose and $Xer_{baseline}$. Step-wise forward selection was based on the largest significant log-likelihood differences

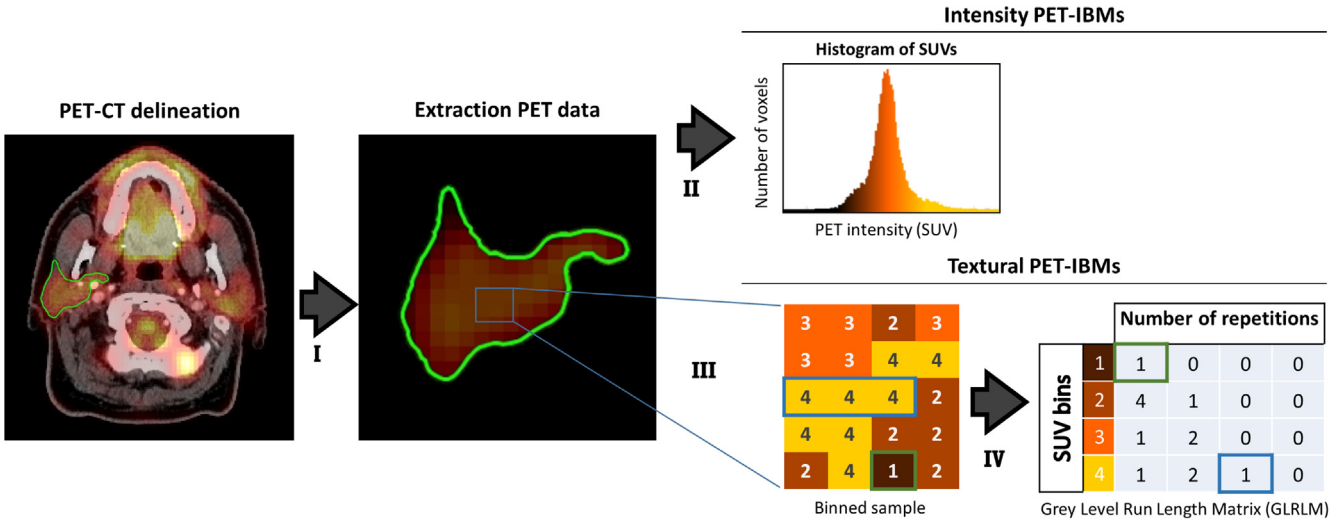

Fig. 1. Example of PET-IBM extraction process. The PET information from the delineated parotid gland was extracted (I). Intensity PET-IBMs were obtained from all voxels inside this contour (II). The SUVs were binned for the textural analysis (III). For illustration, a Grey Level Run length Matrix is constructed from a binned sample, it quantifies the number of repetitions of binned SUVs from left to right (IV).

[27]. Lasso regularisation uses the penalisation term lambda, which excludes variables by reducing their coefficients to zero. The optimal lambda was determined by 100-times repeated 10-fold cross validation [28].

To understand the contribution of the different types of PET-IBMs to the reference model, the model analysis of all SUV intensity and textural PET-IBMs were conducted separately. Subsequently, the resulting SUV intensity and textural models were compared to the reference and the 'mean SUV' model. The performance of the constructed models was quantified with the Area Under the ROC curve (AUC), the Nagelkerke R^2 and the discrimination slope. Furthermore, calibration was evaluated with the Hosmer–Lemeshow test. Internal validation was performed with bootstrapping to correct for optimism of the model [29,30]. Analyses were performed with the R-packages 'Lasso and Elastic-Net Regularized Generalized Linear Models' (version 2.0-2) [28] and 'Regression Modeling Strategies' (version 4.3-1) [31].

Inter-variable relationships

The relationship between variables of predictive PET-IBMs (and $Xer_{baseline}$) was investigated with Pearson correlation (continuous variables) and univariable logistic regression analysis (binary variables). Furthermore, in a previous study, the short run emphasis (SRE), which was extracted from CT information of the parotid gland, was significantly associated with Xer_{12m} [13]. In the current study, this SRE was also extracted from the CT-scans of patients without metal artefacts in the images. Subsequently the correlation of the CT-based SRE values and the predictive PET-IBMs was tested. Additionally, the improvement of the PET-IBM or reference models by SRE was also tested in this patient subset.

Results

Patients

Patient characteristics are depicted in Table 1. Briefly, nearly all patients were bi-laterally irradiated, most patients had oropharyngeal carcinomas and had no baseline xerostomia (none vs. any: 61% vs. 39%). Sixty of the 161 (37%) patients developed moderate-to-severe xerostomia (Xer_{12m}).

Table 1
Patient characteristics.

Characteristics	*N* = 161	%
Sex		
Female	50	31
Male	111	69
Age		
18–65 years	95	59
>65 years	66	41
Tumour site		
Oropharynx	78	48
Nasopharynx	7	4
Hypopharynx	18	11
Larynx	51	32
Oral cavity	7	4
Tumour classification		
T1	14	9
T2	51	32
T3	52	32
T4	44	27
Node classification		
N0	71	44
N1	14	9
N2abc	74	46
N3	2	1
Systemic treatment		
Yes	71	44
No	90	56
Treatment technique		
IMRT	145	90
VMAT	16	10
Bi-lateral		
Yes	139	86
No	22	14
Baseline Xerostomia		
No	98	61
A bit	46	29
Quite a bit	13	8
A lot	4	2

Abbreviations: IMRT: intensity-modulated radiation therapy; VMAT: volumetric arc therapy.

Univariable analysis

In the univariable analysis, the mean dose to the parotid gland and $Xer_{baseline}$ were associated with Xer_{12m}. Univariable analysis

Table 2
Estimated coefficients (uncorrected and corrected for optimism) of reference model and PET-IBM models.

	β		OR (95% CI)	*p*-Value
	Uncorrected	Corrected		
Intercept	−2.633	−2.579		
$Xer_{baseline}$	1.559	1.526	*4.75* (2.32–9.75)	<0.001
PG dose	0.056	0.054	*1.06* (1.02–1.10)	0.002
Intercept	0.906	0.828		
$Xer_{baseline}$	1.473	1.384	4.36 (2.08–9.14)	<0.001
PG dose	0.051	0.047	1.05 (1.01–1.09)	0.007
Mean SUV	**−1.776**	**−1.669**	**0.17 (0.05–0.64)**	**0.009**
Intercept	1.070	0.984		
$Xer_{baseline}$	1.487	1.402	4.43 (2.10–9.31)	<0.001
PG dose	0.050	0.048	1.05 (1.01–1.09)	0.007
P90	**−1.620**	**−1.527**	**0.20 (0.06–0.63)**	**0.006**
Intercept	−2.752	−2.598		
$Xer_{baseline}$	1.577	1.479	4.84 (2.29–10.22)	<0.001
PG dose	0.055	0.051	1.05 (1.02–1.10)	0.004
LRHG3E	**−0.938**	**−0.880**	**0.39 (0.19–0.82)**	**0.013**

Abbreviations: $Xer_{baseline}$: xerostomia at baseline; PG dose: contralateral mean dose to parotid gland; P90: 90th percentile of intensities; LRHG3E: Long Run High Grey-level Emphasis 3; β: regression coefficients; OR: odds ratio; CI: confidence interval.

showed that 11 of 24 intensity PET-IBMs and 35 of 66 textural PET-IBMs were significantly associated with Xer_{12m} (Supplementary data 3). In general, a negative coefficient was observed for PET-IBMs that have a positive relationship with SUVs in the parotid gland, indicating that low parotid gland SUVs were associated with a high Xer_{12m} risk.

Multivariable analysis

Reference model

The reference model with the variables contra-lateral parotid gland dose and $Xer_{baseline}$ (none vs. any) was fit to the dataset (Table 2). The performance measures are depicted in Table 3 (AUC = 0.73 (0.65–0.81), R^2 = 0.22).

Intensity PET-IBMs

First, the basic PET-IBM model ('mean SUV', parotid dose, $Xer_{baseline}$) showed that the addition of the 'mean SUV' significantly improved the reference model (Likelihood ratio test; $p = 0.005$). Consistent with the univariable analysis, the negative regression coefficient of the mean SUV indicates that high mean SUVs were associated with a lower Xer_{12m} risk (Table 2). The performance of this basic PET-IBM model (AUC = 0.77 (0.69–0.84), R^2 = 0.27), was better than that of the reference model (Table 3).

Resulting from both the Lasso regularisation and forward selection, the 90th percentile of SUVs (P90) was the most predictive of all intensity PET-IBMs (Fig. 2), leading to a significant (Likelihood-ratio test; $p = 0.002$), substantial improvement of the model performance measures (Tables 2 and 3; AUC = 0.77 (0.70–0.84), R^2 = 0.28) compared to the reference model (AUC = 0.73 (0.64–0.83), R^2 = 0.23). High correlations were observed between P90 and the IBMs that could also significantly improve the reference model when individually added to the reference model ($\rho = 0.82 \pm 0.15$). See Supplementary data 4 for the correlations of PET-IBMs.

In Fig. 3 the NTCP curves for different P90 values are depicted of the following P90 model:

$$\mathrm{NTCP} = \frac{1}{1 - e^{-s}}$$
$$\text{where } s = 0.984 + 0.048 \cdot \text{Contra Dose (PG)} + 1.402 \cdot Xer_{baseline} - 1.527 \cdot \text{P90(PG)}$$

Table 3
Performance of NTCP models with and without PET-IBMs.

		Reference model		PET-IBM models	
		$Xer_{baseline}$ PG dose –	$Xer_{baseline}$ PG dose mean SUV	$Xer_{baseline}$ PG dose P90	$Xer_{baseline}$ PG dose LRHG3E
Overall	−2 log-likelihood Nagelkerke R^2	184.51 0.22	176.57 0.27	175.30 0.28	174.31 0.29
Discrimination	Area Under the Curve (AUC) Discrimination slope	0.73 (0.65–0.81) 0.17	0.77 (0.69–0.84) 0.20	0.77 (0.70–0.84) 0.21	0.77 (0.69–0.84) 0.21
Calibration	HL test X^2 (*p*-value) Calibration slope (intercept)	11.22 (0.19) 1.00 (0.00)	4.24 (0.83) 0.95 (0.02)	6.72 (0.57) 0.95 (0.02)	6.30 (0.61) 0.99 (0.00)
Internal validation	$AUC_{corrected}$ Nagelkerke $R^2_{corrected}$	0.72 0.20	0.75 0.24	0.76 0.25	0.75 0.26

Abbreviations: HL: Hosmer–Lemeshow; corrected: corrected for optimism with bootstrapping; IBM: Image Biomarker; $Xer_{baseline}$: xerostomia at baseline; PG dose: contralateral mean dose to parotid gland; P90: 90th percentile of intensities; LRHG3E: Long Run High Grey-level Emphasis 3.

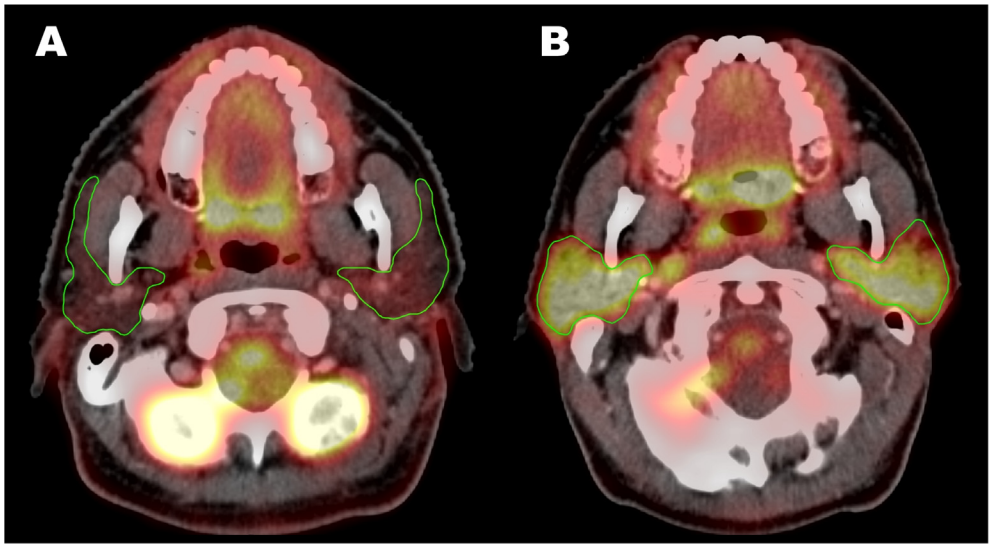

Fig. 2. Example of patients with (A) low and (B) high values of mean SUV, P90 and LRHG3E, which were associated with (A) higher and (B) lower risk of developing Xer_{12m}. Scaling in both images: 0.5–3.5 SUV.

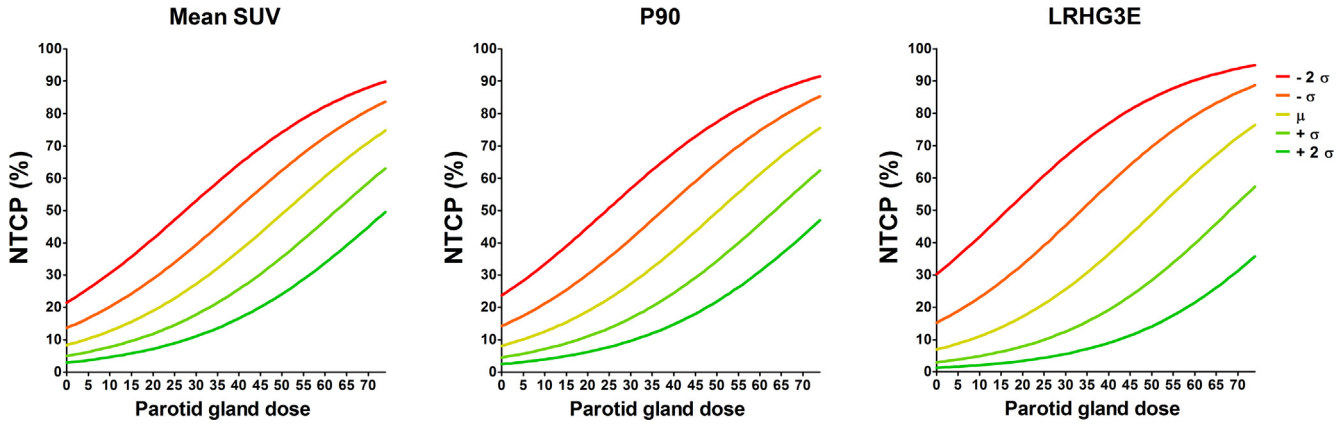

Fig. 3. Normal Tissue Complication Probability (NTCP) values for late xerostomia (Xer_{12m}) of models based on mean SUV (left), P90 (middle) and LRHG3E (right). Curves are given for the mean PET-IBM values (P90: $\mu = 2.23$; LRHG3E: $\mu = 201.24$) and for 1 and 2 standard deviation from these mean values (mean SUV: $\mu = 1.93$, $\sigma = 0.33$; P90: $\mu = 2.23$, $\sigma = 0.41$; LRHG3E: $\mu = 201.24$, $\sigma = 177.05$). For these curves no baseline xerostomia was assumed ($Xer_{baseline} = 0$).

Textural PET-IBMs

The most predictive textural PET-IBM was the Long Run High Grey-level Emphasis 3 (LRHG3E), which is derived from the GLRLM. The value of this PET-IMB increases when long repetitions of high SUVs are present in the parotid gland with extra (power of 3) emphasis on high SUVs (see Supplementary data 2 for formula). This variable was selected by both the Lasso regularisation and the step-wise forward selection. This variable significantly improved the reference model in predicting Xer_{12m} (Likelihood-ratio test; $p = 0.001$). The negative coefficient of LRHG3E indicated once more that high SUVs are associated with low Xer_{12m} risk (Table 2). The addition of LRHG3E improved the reference model performance (0.77 (0.69–0.84), $R^2 = 0.29$; Table 3). The NTCP curves for different LRHG3E are depicted in Fig. 3 for the following model:

$$\mathrm{NTCP} = \frac{1}{1 - e^{-s}}$$

$$\text{where } s = -2.598 + 0.051 \cdot \text{Contra Dose (PG)} + 1.479 \cdot \mathrm{Xer}_{\mathrm{baseline}} - 0.880 \cdot \frac{\mathrm{LRHG3E(PG)} - 201.24}{177.05}$$

Inter-variable relationships

The predictive PET-IBM P90 (intensity) and LRHG3E (textural) were closely correlated ($p < 0.001$; $r = 0.83$). Moreover, they did not add independent information to each other in predicting Xer_{12m} (Likelihood ratio test; $p > 0.21$). Univariable logistic analysis showed no significant association between $Xer_{baseline}$ and P90 ($p = 0.079$) or LRHG3E ($p = 0.465$).

In the current study cohort, 100 patients did not have metal artefacts in the CT images and could therefore be used for the anal-

ysis of the CT-based IBM, the short run emphasis (SRE) [13]. This CT-based SRE was significantly correlated to the predictive PET-IBM P90 ($p = 0.008$; $r = -0.26$) and LRHG3E ($p = 0.026$; $r = -0.22$). The SRE neither significantly improved the reference model (likelihood ratio test, $p = 0.055$), nor did it add to the PET-IBM models with P90 (likelihood ratio test, $p = 0.140$) and LRHG3E (likelihood ratio test, $p = 0.096$) in this cohort subset.

Discussion

This study is novel to show that the high metabolic activity of the parotid gland was associated with a lower risk of developing late xerostomia (Xer_{12m}). Moreover, the prediction of late xerostomia was significantly and substantially improved with addition of patient-specific PET-IBMs to the reference model based on dose and $Xer_{baseline}$. These findings could improve understanding of normal tissue response following radiotherapy, since the variation in patient-specific PET characteristics can partly explain the unexplained variance in predicting xerostomia with dose parameters. Moreover, it could improve identification of patients that are at risk of late radiation-induced side effects, which could potentially benefit most from new therapy technology such as proton [7] and MRI-guided irradiation [8]. In other words, better prediction of toxicities could improve the treatment decision support [9,32]. However, external validation of the PET-IBM models in an independent dataset is necessary before clinical implementation [33].

The PET-IBM that indicates the minimum value of the 90% highest SUVs (P90) was the most predictive of all intensity PET-IBMs. The mean SUV also performed well, but P90 appeared more relevant in this dataset. A high P90 was associated with a lower risk of developing late xerostomia. Similar effect and predictive improvement was observed from LRHG3E (Long Run High Grey-level Emphasis 3) of the textural PET-IBMs, which significantly correlated with P90 ($\rho = 0.83$). This PET-IBM indicates high SUVs that are spatially adjacent to each other. Both PET-IBMs were negatively associated with Xer_{12m}, suggesting that patients with low metabolic activity in the parotid glands were at risk of developing late xerostomia. Although both P90 and LRHG3E perform similarly, currently the P90 is simpler to calculate. However, LRHG3E also contains information about the spatial connectivity of the high SUV voxels, i.e. large repetitions of voxels with high SUV increase the LRHG3E values. External validation is needed to confirm the predictive power of LRHG3E over P90. Additionally, an alternative variable selection approach, Lasso regularisation, resulted in very comparable final models. Since they were independent of the method of analysis, it suggests that the associations in this dataset were relatively stable.

Predictive PET-IBMs were not significantly associated with $Xer_{baseline}$. This suggests that PET-IBMs contain unique and additional information to baseline xerostomia complaints, since the addition of PET-IBMs to $Xer_{baseline}$ (and PG dose) improved the prediction of Xer_{12m} significantly.

This study suggests that high metabolic parotid glands have more viable cells (parenchyma and/or stem cells) with more repair capability and/or are less radiosensitive. Although possibly driven by multiple underlying biological processes, there is some similarity in the tumour reaction to radiation. For tumour tissue it is known that high metabolic tumours are more likely to recur [34], particularly in their high metabolic regions [35]. A possible explanation is that it arises from a combination of higher cell density, proliferation rate of metabolically active tissue and DNA repair capacity [36].

Other studies have shown that parotid gland SUVs decrease post-radiotherapy, and in addition that this change was associated with parotid gland dose [37,38]. Cannon et al. [38] showed that mean 'SUV-weighted parotid gland dose (voxel-wise)' was significantly related to fractional-SUV (post-SUV/pre-SUV). In an additional small cohort ($n = 8$), they showed that fractional-SUV was significantly associated with fractional salivary flow and physician-rated xerostomia. Although this indirectly suggests that 'SUV-weighted parotid gland dose' is related to xerostomia measures, the direct and separate associations of parotid gland dose and pre-treatment SUV with xerostomia measures or fractional SUV were unfortunately not described.

In previous work, a positive association was shown between higher risk of developing late xerostomia and CT-based SRE (Short Run Emphasis), which might be related to the ratio between non-functional fatty tissue and functional parotid parenchyma tissue. In this study, we showed that this CT-IBM was significantly correlated to P90 and LRGH3E in patients without metal artefacts ($n = 100$) and did not significantly add to the PET-IBM models. Additionally, the performance of predicting Xer_{12m} was substantially higher with PET-based IBM models than with than CT-based IBMs. This suggests that ^{18}F-FDG PET is better to quantify the ratio between fatty non-functional and functional parotid parenchyma tissue. This is logical since ^{18}F-FDG PET is a functional image modality. Furthermore, the SRE did not show a significant improvement in the reference model for the cohort subset, which might be caused by the small additive effect of SRE and low number of patients on which this IBM could be tested.

A well-defined protocol was used to ensure optimal standardisation of SUV in the ^{18}F-FDG PET images by correcting for body-weight, injection dose, tracer uptake period, and glucose plasma levels by letting the patients fast [19,20]. Although SUVs may also be affected by fasting blood glucose level, muscle activity, liver and kidney function, the images were not corrected for these fluctuations. Furthermore, patients with metal artefacts in CT images were included, where the attenuation correction can influence SUVs, but this bias will primarily be located around the metal implant [20]. Additional analyses showed that the PET-IBMs' performance was still good in the sub cohort of patients without metal artefacts. Additionally, future improvements of the consistency and spatial resolution of PET imaging should also improve the performance of the PET-IBMs in predicting Xer_{12m}.

In this study, patient-rated outcomes (EORTC QLQ-H&N35 questionnaire) were used as a measure for moderate-to-severe xerostomia, because of their relationship with the quality of life of HNC patients [10]. However, some unexplained variability of the models may be caused by the assessment of xerostomia, as the questionnaires can be interpreted differently by the individual patients [39]. Our current study could be strengthened by the addition of investigating the associations between PET-IBMs and objective xerostomia measures. Parotid flow rates are often used, but several studies have shown no or modest correlation between patient reported xerostomia and parotid flow rates [40] and have a low reproducibility [41]. Another example is scintigraphy of parotid gland ejection fraction over time. Although this technique seems promising as a quantitative measure for xerostomia, it requires additional scans with complex procedures with radioactive tracers [41]. This highlights the importance for future research on a non-invasive, accessible and reliable quantitative measure of xerostomia. Nevertheless, we believe that patient-rated xerostomia remains an important endpoint, due to its clinical importance and practical benefits.

Conclusion

The pre-treatment PET-IBMs indicated that a large quantity of high SUVs in the parotid gland was significantly associated with a lower risk of developing xerostomia 12 months after radiotherapy. The addition of the predictive intensity PET-IBM

(90th percentile of SUV) to a model with parotid gland dose and baseline xerostomia improved the prediction performance of the reference model substantially (from 0.73 (0.65–0.81) to 0.77 (0.70–0.84)). This study highlights the importance of incorporating patient-specific functional characteristics into NTCP model development and can, thereby, contribute to the understanding of the patient-specific response of healthy tissue to radiation dose.

Conflict of interest

The authors state that the research presented in this manuscript is free of conflicts of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.radonc.2017.08.024.

References

[1] Leclerc M, Lartigau E, Lacornerie T, Daisne JF, Kramar A, Gregoire V. Primary tumor delineation based on (18)FDG PET for locally advanced head and neck cancer treated by chemo-radiotherapy. Radiother Oncol 2015;116:87–93.

[2] Stuckensen T, Kovács AF, Adams S, Baum RP. Staging of the neck in patients with oral cavity squamous cell carcinomas: a prospective comparison of PET, ultrasound, CT and MRI. J Cranio-Maxillofacial Surg 2000;28:319–24.

[3] Abgral R, Keromnes N, Robin P, Le Roux P-Y, Bourhis D, Palard X, et al. Prognostic value of volumetric parameters measured by (18)F-FDG PET/CT in patients with head and neck squamous cell carcinoma. Eur J Nucl Med Mol Imaging 2014;41:659–67.

[4] Eary JF, O'Sullivan F, O'Sullivan J, Conrad EU. Spatial heterogeneity in Sarcoma ^{18}F-FDG uptake as a predictor of patient outcome. J Nucl Med 2008;49:1973–9.

[5] Koyasu S, Nakamoto Y, Kikuchi M, Suzuki K, Hayashida K, Itoh K, et al. Prognostic value of pretreatment ^{18}F-FDG PET/CT parameters including visual evaluation in patients with head and neck squamous cell carcinoma. AJR Am J Roentgenol 2014;202:851–8.

[6] Alluri KC, Tahari AK, Wahl RL, Koch W, Chung CH, Subramaniam RM. Prognostic value of FDG PET metabolic tumor volume in human papillomavirus-positive stage III and IV oropharyngeal squamous cell carcinoma. AJR Am J Roentgenol 2014;203:897–903.

[7] Lomax A. Intensity modulation methods for proton radiotherapy. Phys Med Biol 1999;44:185–205.

[8] Lagendijk JJW, Raaymakers BW, Raaijmakers AJE, Overweg J, Brown KJ, Kerkhof EM, et al. MRI/linac integration. Radiother Oncol 2008;86:25–9.

[9] Langendijk JA, Lambin P, De Ruysscher D, Widder J, Bos M, Verheij M. Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach. Radiother Oncol 2013;107:267–73.

[10] Langendijk JA, Doornaert P, Verdonck-de Leeuw IM, Leemans CR, Aaronson NK, Slotman BJ. Impact of late treatment-related toxicity on quality of life among patients with head and neck cancer treated with radiotherapy. J Clin Oncol 2008;26:3770–6.

[11] Beetz I, Schilstra C, Van Der Schaaf A, Van Den Heuvel ER, Doornaert P, Van Luijk P, et al. NTCP models for patient-rated xerostomia and sticky saliva after treatment with intensity modulated radiotherapy for head and neck cancer: the role of dosimetric and clinical factors. Radiother Oncol 2012;105:101–6.

[12] Vergeer MR, Doornaert PAH, Rietveld DHF, Leemans CR, Slotman BJ, Langendijk JA. Intensity-modulated radiotherapy reduces radiation-induced morbidity and improves health-related quality of life: results of a nonrandomized prospective study using a standardized follow-up program. Int J Radiat Oncol Biol Phys 2009;74:1–8.

[13] van Dijk LV, Brouwer CL, van der Schaaf A, Burgerhof JGM, Beukinga RJ, Langendijk JA, et al. CT image biomarkers to improve patient-specific prediction of radiation-induced xerostomia and sticky saliva. Radiother Oncol 2017;122:185–91.

[14] Van Der Laan HP, Gawryszuk A, Christianen MEMC, Steenbakkers RJHM, Korevaar EW, Chouvalova O, et al. Swallowing-sparing intensity-modulated radiotherapy for head and neck cancer patients: Treatment planning optimization and clinical introduction. Radiother Oncol 2013;107:282–7.

[15] Christianen MEMC, Langendijk JA, Westerlaan HE, Van De Water TA, Bijl HP. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. Radiother Oncol 2011;101:394–402.

[16] Beetz I, Schilstra C, Burlage FR, Koken PW, Doornaert P, Bijl HP, et al. Development of NTCP models for head and neck cancer patients treated with three-dimensional conformal radiotherapy for xerostomia and sticky saliva: the role of dosimetric and clinical factors. Radiother Oncol 2012;105:86–93.

[17] Houweling AC, Philippens MEP, Dijkema T, Roesink JM, Terhaard CHJ, Schilstra C, et al. A comparison of dose-response models for the parotid gland in a large group of head-and-neck cancer patients. Int J Radiat Oncol Biol Phys 2010;76:1259–65.

[18] Jellema AP, Doornaert P, Slotman BJ, Leemans CR, Langendijk JA. Does radiation dose to the salivary glands and oral cavity predict patient-rated xerostomia and sticky saliva in head and neck cancer patients treated with curative radiotherapy? Radiother Oncol 2005;77:164–71.

[19] Boellaard R, Oyen WJG, Hoekstra CJ, Hoekstra OS, Visser EP, Willemsen AT, et al. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. Eur J Nucl Med Mol Imaging 2008;35:2320–33.

[20] Boellaard R, Delgado-Bolton R, Oyen WJG, Giammarile F, Tatsch K, Eschner W, et al. FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0. Eur J Nucl Med Mol Imaging 2014;42:328–54.

[21] Haralick R, Shanmugan K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3:610–21.

[22] Tang X. Texture information in run-length matrices. IEEE Trans Image Process 1998;7:1602–9.

[23] Galloway MM. Texture analysis using gray level run lengths. Comput Graph Image Process 1975;4:172–9.

[24] Thibault G, Fertil B, Navarro C, Pereira S, Cau P, Levy N, et al. texture indexes and gray level size zone matrix application to cell nuclei classification. Pattern Recognit Inf Process 2009:140–5.

[25] Amadasun M, King R. Textural features corresponding to textural properties. IEEE Trans Syst Man Cybern 1989;19:1264–73.

[26] Leijenaar RTH, Nalbantov G, Carvalho S, van Elmpt WJC, Troost EGC, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. Sci Rep 2015;5:11075.

[27] Dehing-Oberije C, De Ruysscher D, Petit S, Van Meerbeeck J, Vandecasteele K, De Neve W, et al. Development, external validation and clinical usefulness of a practical prediction model for radiation-induced dysphagia in lung cancer patients. Radiother Oncol 2010;97:455–61.

[28] Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 2010;33.

[29] Van Der Schaaf A, Xu CJ, Van P, Luijk, Van't AA, Veld, Langendijk JA, Schilstra C. Multivariate modeling of complications with data driven variable selection: Guarding against overfitting and effects of data set size. Radiother Oncol 2012;105:115–21.

[30] Harrell FE. Regression modeling strategies. With applications to linear models, logistic regression, and survival analysis. 2001.

[31] R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: the R Foundation for Statistical Computing. 2011: Available online at http://www.R-project.org/.

[32] Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. Magn Reson Imaging 2012;30:1234–48.

[33] Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. Eur Urol 2015;67:1142–51.

[34] Xie P, Yue J-B, Fu Z, Feng R, Yu J-M. Prognostic value of ^{18}F-FDG PET/CT before and after radiotherapy for locally advanced nasopharyngeal carcinoma. Ann Oncol 2010;21:1078–82.

[35] Due AK, Vogelius IR, Aznar MC, Bentzen SM, Berthelsen AK, Korreman SS, et al. Recurrences after intensity modulated radiotherapy for head and neck squamous cell carcinoma more likely to originate from regions with high baseline [^{18}F]-FDG uptake. Radiother Oncol 2014;111:360–5.

[36] Jeong J, Deasy JO. Modeling the relationship between fluorodeoxyglucose uptake and tumor radioresistance as a function of the tumor microenvironment. Comput Math Methods Med 2014;2014:847162.

[37] Roach MC, Turkington TG, Higgins KA, Hawk TC, Hoang JK, Brizel DM. FDG-PET assessment of the effect of head and neck radiotherapy on parotid gland glucose metabolism. Int J Radiat Oncol Biol Phys 2012;82:321–6.

[38] Cannon B, Schwartz DL, Dong L. Metabolic imaging biomarkers of postradiotherapy xerostomia. Int J Radiat Oncol Biol Phys 2012;83:1609–16.

[39] Meirovitz A, Murdoch-Kinch CA, Schipper M, Pan C, Eisbruch A. Grading xerostomia by physicians or by patients after intensity-modulated radiotherapy of head-and-neck cancer. Int J Radiat Oncol Biol Phys 2006;66:445–53.

[40] Eisbruch A, Kim HM, Terrell JE, Marsh LH, Dawson LA, Ship JA. Xerostomia and its predictors following parotid-sparing irradiation of head-and-neck cancer. Int J Radiat Oncol Biol Phys 2001;50:695–704.

[41] Cheng SCH, Wu VWC, Kwong DLW, Ying MTC. Assessment of post-radiotherapy salivary glands. Br J Radiol 2011;84:393–402.